UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/006,984	12/04/2001	Jan K. Skoglund	020184-000510US	1945

20350 7590 10/04/2004

TOWNSEND AND TOWNSEND AND CREW, LLP
TWO EMBARCADERO CENTER
EIGHTH FLOOR
SAN FRANCISCO, CA 94111-3834

EXAMINER	
ALBERTALLI, BRIAN LOUIS	
ART UNIT	**PAPER NUMBER**
2655	

DATE MAILED: 10/04/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/006,984	SKOGLUND ET AL.	
	Examiner	Art Unit	
	Brian L Albertalli	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) 1-21 is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) 1-21 is/are rejected.

7)☒ Claim(s) 5 is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on 12 February 2002 is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Specification

The title of the invention is not descriptive. A new title is required that is clearly indicative of the invention to which the claims are directed.

Claim Rejections - 35 USC § 112

1. The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

The term "long sequence of previous frames" in claim 4 is a relative term which renders the claim indefinite. The term "long sequence of previous frames" is not defined by the claim, the specification does not provide a standard for ascertaining the requisite degree, and one of ordinary skill in the art would not be reasonably apprised of the scope of the invention. The specification only mentions that a "sufficiently long window" should be used for long term correction (page 10, paragraph 47, lines 1-2). It is not clear whether a "sufficiently long window" is on the order of a few frames, several seconds of speech, or several minutes of speech. Accordingly, the term "long sequence of previous frames" has been interpreted herein to encompass any number of previous frames.

Claim Rejections - 35 USC § 103

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

3. Claims 1-8, 10-11, 13-15, and 18 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sohn et al. (*A Voice Activity Detector Employing Soft Decision Based Noise Spectrum Adaptation*).

In regard to claim 1, Sohn et al. discloses a method for detecting speech activity for a signal, the method comprising the steps of:

extracting a plurality of features from the signal (DFT coefficients, page 365, second column, section 2, lines 15-22);

modeling a first and a second probability density functions (PDFs) of the plurality of features, wherein:

the first PDF models active speech conditions for the signal (equation 5, probability of noisy speech X given speech is present H1), and

the second PDF models inactive speech conditions for the signal (equation 4, probability of a noisy speech X, given speech is absent H0);

adapting the second PDF to respond to changes in the signal over time (the noise spectrum is continuously updated, equation 16 and page 368, first column, lines 6-8);

probability-based classifying of the signal based, at least in part, on the plurality of features (decision rule is used to differentiate between silence and noise, equation 7); and

distinguishing speech in the signal based, at least in part, upon the probability-based classifying step (the decision rule is a decision whether speech is present H1, or absent H0, see page 366, first column, lines 11-12).

Sohn et al. does not disclose that the speech PDF is adapted.

Official notice is taken that it is notoriously well known and recognized in the art that speech signals are nonstationary, that is, their statistical models change with respect to time.

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Sohn et al. to adapt the speech PDF as well as the noise PDF, since the statistics of the speech would change over time. This would make the speech PDF model the actual speech more accurately, which would increase the probability of correct speech/non-speech decisions.

In regard to claim 2, Sohn et al. discloses the probability based classifying step uses first and second PDFs (equation 7, classification decision is dependent on PDFs given in equations 4 and 5).

In regard to claim 3, Sohn et al. discloses the modeling step comprises a step of determining a mathematical model (PDFs) for the signal from the plurality of features (the variances of the noise and speech are determined from the power spectra of the noise, equations 1 and 2; which are used to determine the PDFs for the signals, page 365, second column, section 2 line 15 through page 366, equation 2).

In regard to claim 4, Sohn et al. discloses the adapting step comprises increasing a likelihood (equation 16, the noise model converges towards the actual noise, page 368, first column, lines 2-4).

In regard to claim 5, Sohn et al. discloses adapting step comprises a step of identifying extreme values in a long sequence of previous frames (page 367, adaptation formula is a recursive formula based on the current frame m and a previous frame m-1, equation 16, and second column, fourth paragraph).

In regard to claim 6, Sohn et al. discloses the probability-based classifying step comprises a step of classifying based on likelihood ratio detection (a log likelihood ratio is used for the decision rule, page 366, equation 7).

In regard to claim 7, Sohn et al. discloses the probability-based classifying step comprises applying a log-likelihood ratio test to one of the plurality of features (page

366, equation 7, the log likelihood ratio is based on the variances of the speech and noise, which are determined from the coefficients from the DFT, page 365, second column, section 2 line 15 through page 366, equation 2).

In regard to claim 10, Sohn et al. discloses at least one of the first and second PDFs comprises a plurality of basic density models (page 366, equations 4 and 5, each PDF is the product of L basic density models).

In regard to claim 11, Sohn et al. discloses at least one of the plurality of features is related to power in a spectral band of the signal (DFT coefficients are determined, the coefficients denote the true power spectra of the noise and speech, page 366, first column, line 3).

In regard to claims 13 and 18, Sohn et al. does not explicitly disclose a computer-readable medium having computer-executable instructions for performing the computer-implementable method for detecting speech activity for the signal of claim 1 or 14.

Official notice is taken that it is notoriously well recognized to implement a signal processing method on a computer and to store instructions for implementing the method on a computer readable medium.

It would have been obvious to one of ordinary skill in the art at the time of invention to store the method as disclosed by Sohn et al. on as computer readable code on a computer readable medium, so the method could be implemented on computer.

In regard to claim 14, Sohn et al. discloses a method for detecting sound activity for a signal, the method comprising the steps of:

extracting a plurality of features from the signal (DFT coefficients, page 365, second column, section 2, lines 15-22);

modeling an active speech probability density function (PDF) of the plurality of features (equation 5, probability of noisy speech X given speech is present H1);

modeling an inactive speech PDF of the plurality of features (equation 4, probability of a noisy speech X, given speech is absent H0);

adapting the inactive speech PDFs to respond to changes in the signal over time (the noise spectrum is continuously updated, equation 16 and page 368, first column, lines 6-8);

probability-based classifying of the signal based, at least in part, on the plurality of features (decision rule is used to differentiate between silence and noise, equation 7); and

distinguishing speech in the signal based, at least in part, upon the probability-based classifying step (the decision rule is a decision whether speech is present H1, or absent H0, see page 366, first column, lines 11-12).

Sohn et al. does not disclose that the speech PDF is adapted.

Official notice is taken that it is notoriously well known and recognized in the art that speech signals are nonstationary, that is, their statistical models change with respect to time.

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Sohn et al. to adapt the speech PDF as well as the noise PDF, since the statistics of the speech would change over time. This would make the speech PDF model the actual speech more accurately, which would increase the probability of correct speech/non-speech decisions.

In regard to claim 15, Sohn et al. discloses the probability-based classifying step uses the active and inactive speech PDFs (equation 7, classification decision is dependent on PDFs given in equations 4 and 5).

In regard to claim 16, Sohn et al. discloses the adapting step comprises a step of increasing a likelihood (equation 16, the noise model converges towards the actual noise, page 368, first column, lines 2-4).

4. Claim 8 is rejected under 35 U.S.C. 103(a) as being unpatentable over Sohn et al., in view of Huang et al. (U.S. Patent 6,421,641).

Sohn et al. does not disclose at least one of the first and second PDFs comprises a Gaussian mixture model (page 365, second column section 2, lines 15-19).

Huang et al. discloses an adaptable method of modeling features of a speech signal as a Gaussian mixture model (Fig. 3, column 5, lines 42-44).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Sohn et al. to model the features of the speech as a Gaussian mixture model, since the model disclosed by Huang et al. provides a real-time adaptation which would speed the entire classification process, thereby reducing the need for subsequent hangover correction.

5. Claims 9, 17, and 19-21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sohn et al., in view of Paez et al. (*Minimum Mean Squared Error Quantization in Speech PCM and DPCM Systems*).

In regard to claims 9 and 17, Sohn et al. does not disclose at least one of the first and second PDFs comprises a non-Gaussian model.

Paez et al. discloses that speech most closely approximates a gamma probability density function (page 227, first column, first and second paragraphs, and page 226, Fig. 3).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Sohn et al. to use a non-Gaussian model for one of the PDFs, since a non-Gaussian model would model the actual speech more accurately, which would increase the probability of correct speech/non-speech decisions.

In regard to claim 19, Sohn et al. discloses a method for detecting sound activity for a signal, the method comprising the steps of:

extracting a plurality of features from the signal (DFT coefficients, page 365, second column, section 2, lines 15-22);

modeling an active speech probability density function (PDF) of the plurality of features (equation 5, probability of noisy speech X given speech is present H1);

modeling an inactive speech PDF of the plurality of features (equation 4, probability of a noisy speech X, given speech is absent H0);

adapting the inactive speech PDFs to respond to changes in the signal over time (the noise spectrum is continuously updated, equation 16 and page 368, first column, lines 6-8);

probability-based classifying of the signal based, at least in part, on the plurality of features (decision rule is used to differentiate between silence and noise, equation 7); and

distinguishing speech in the signal based, at least in part, upon the probability-based classifying step (the decision rule is a decision whether speech is present H1, or absent H0, see page 366, first column, lines 11-12).

Sohn et al. does not disclose that the speech PDF is adapted.

Official notice is taken that it is notoriously well known and recognized in the art that speech signals are nonstationary, that is, their statistical models change with respect to time.

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Sohn et al. to adapt the speech PDF as well as the noise PDF, since the statistics of the speech would change over time. This would make the speech PDF

model the actual speech more accurately, which would increase the probability of correct speech/non-speech decisions.

Furthermore, Sohn et al. does not disclose that at least one of the active and inactive speech PDFs uses a non-Gaussian model.

Paez et al. discloses that speech most closely approximates a gamma probability density function (page 227, first column, first and second paragraphs, and page 226, Fig. 3).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Sohn et al. to use a non-Gaussian model for one of the PDFs, since a non-Gaussian model would model the actual speech more accurately, which would increase the probability of correct speech/non-speech decisions.

In regard to claim 20, the combination of Sohn et al. and Paez et al., as applied to claim 19, above, discloses modeling the active speech as a non-Gaussian model.

Neither Sohn et al. nor Paez et al. disclose modeling the inactive speech as a non-Gaussian model.

Official notice is taken that it is notoriously well known and recognized in the art that inactive speech (background noise) can be modeled by non-Gaussian models.

It would have been obvious to one of ordinary skill in the art at the time of invention to model both the active and inactive speech PDFs using a non-Gaussian model, since a non-Gaussian model would model the actual speech and the inactive

speech more accurately, which would increase the probability of correct speech/non-speech decisions.

In regard to claim 21, Sohn et al. and Paez et a. do not explicitly disclose a computer-readable medium having computer-executable instructions for performing the computer-implementable method for detecting speech activity for the signal of claim 19.

Official notice is taken that it is notoriously well recognized to implement a signal processing method on a computer and to store instructions for implementing the method on a computer readable medium.

It would have been obvious to one of ordinary skill in the art at the time of invention to store the method as disclosed by Sohn et al. on as computer readable code on a computer readable medium, so the method could be implemented on computer.

6. Claim 12 is rejected under 35 U.S.C. 103(a) as being unpatentable over Sohn et al. (*A Voice Activity Detector Employing Soft Decision Based Noise Spectrum Adaptation*), hereinafter referred to as Sohn 1, in view of Sohn et al. (*A Statistical Model-Based Voice Activity Detection*), hereinafter referred to as Sohn 2.

Sohn 1 does not disclose a step of smoothing an activity decision for hangover periods to produce a smoothed activity decision.

Sohn 2 discloses a step of smoothing an activity decision for hangover periods to produce a smoothed activity decision (a smoothing factor obtained by equation 11 is

used to modify the final decision statistic, page 2, second column, paragraphs three and four).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Sohn 1 to smooth the activity decision for hangover periods, in order to prevent the clipping of weak speech tails, as disclosed by Sohn 2 (page 2, first column, section III, lines 1-3).

Conclusion

7. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Liu et al. (U.S. Patent 6,615,170) discloses a method for voice activity detection based on a log-likelihood ratio. Endo et al. (U.S. Patent 6,490,554) discloses a method for voice activity detection based on a statistical analysis. Krasney et al. (U.S. Patent 6,349,278) discloses voice activity detection based on a soft decision. Anderson et al. (U.S. Patent 6,453,285) discloses a voice activity detector based on statistics of the speech signal. Sato et al. (U.S. Patent 6,044,342) discloses a method of adjusting a voice activity detection threshold based on statistics.

8. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Brian L Albertalli whose telephone number is (703) 305-1817. The examiner can normally be reached on Mon - Fri, 8:00 AM - 5:30 PM, every second Fri off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Smits can be reached on (703) 305-3011. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

BLA 9/27/04

SUSAN MCFADDEN
PRIMARY EXAMINER